Title: **Sensitivity for Correlated Input Variables and Propagated Errors in Evapotranspiration Estimates from a Humid Region**

May 1995

DRE-325

By: **Hosung Ahn**, Ph. D., P.E.

Hydrogeology Division, WRE,

South Florida Water Management District,

West Palm Beach, Florida

Title: Sensitivity for Correlated Input Variables and Propagated Errors in Evapotranspiration Estimates from a Humid Region.

Author: Hosung Ahn, Hydrogeology Div. WRE, SFWMD

Target Journal: Water Resources Research

Summary:

Accurate estimation of evapotranspiration (ET) is critical to the management of water resources and related environmental phenomena. Due to the difficulty and high cost of direct measurement of ET rates, indirect estimates are commonly used in practice. However, the methodology and required data are largely dependent on the temporal and spatial scales of each particular problem. To handle scaling problems in ET estimates, information on sensitivity and uncertainty in estimated ET are requested in advance, which was the main objective in this paper.

The short term ET rates can be accurately estimated by the Penman-Brutsaert model with meteorological data, including net radiation, atmospheric pressure, air temperature, relative humidity, and wind velocity. In data from a weather station at the ENR project site, a strong correlation was found between net radiation and relative humidity, and sensitivity analyses were performed with and without this correlation effect. Then, the sensitivity result of both cases were evaluated by the conditional probabilities of the sample. The results show that the conditional probability with the correlation effect was usually higher than without it, implying that sensitivity with the correlation effect is more suitable for actual meteorological conditions. In addition, the propagated errors in the estimated ET rates caused by the random errors in the measurement of meteorological data were computed. Both sensitivity and error analyses reveal that the net radiation was the most sensitive and apt to cause possible errors in ET estimates due to its significant random error.

The results of this study will help the District to understand uncertainties in estimated ET rates and to improve the monitoring network of meteorological data in the south Florida region. The proposed sensitivity method with correlation effect can be applicable not only to the QA/QC of the field measurement data, but to any type of hydrologic model used in water supply estimations and environmental management.

Abstract.

A sensitivity and uncertainty analysis of the Penman-Brutsaert evapotranspiration model was conducted using the actual meteorological data collected from a experimental weather station located in the humid south Florida region. Since net radiation and relative humidity were found to be highly correlated in this humid region, a method was developed to analyze the sensitivity with the correlation effect of these two independent variables. After conducting sensitivity analyses with and without the correlation effect, the results were evaluated using conditional probability density functions of both cases. Both theoretical and computed results show that the conditional probability with the correlation effect increases in proportion to the increasing absolute correlation coefficient of two independent variables. This finding suggests that the sensitivity with the correlation effect is more suitable for actual meteorological conditions than that without the correlation effect. After defining the random errors in each independent variable, the propagated errors in the estimated evapotranspiration rate due to erroneous independent variables were computed, which are shown to be valuable for interpretation of the estimated evapotranspiration rates and designing a monitoring network for meteorological data in humid region.

1. Introduction

Accurate estimation of evapotranspiration (ET) is a critical issue in south Florida where the ET rate is one of the largest components of the water budget due to the abundance of wetlands and lakes. Since direct measurement of ET is difficult and expensive, indirect estimation methods are preferable in practice. ET can be estimated indirectly by several different approaches, which can be classified into several different categories including energy budget, aerodynamic, combination of energy budget and aerodynamic, water budget, empirical. Overview of these approaches are presented in several literatures [*Jensen et al.*, 1990; *Brutsaert*, 1982; *Stannard*, 1993; *Winter et al.*, 1995; etc.]. Each approach has its own advantages and disadvantages and choosing an appropriate model is largely dependent on the temporal and spatial scales of the individual problem. For a vast wet surface, where the local advection is minimal, a simple energy budget approach such as Penman equation can be used [*Brutsaert*, 1982]. *Jones et al.* [1984] demonstrated that, for the humid climate in Florida, the Penman method is superior to the other tested methods because it is based on physical derivations with less empiricism. Among several Penman type models, the Penman-Brutsaert (P-B) model was selected here, since this model is known to be adequate for the simulation of ET rate at short time intervals [*Brutsaert*, 1979, 1982; *Stricker and Brutsaert*, 1978; *Katul and Parlange*, 1992]. To estimate ET rates by the P-B model, five meteorological variables including net radiation, atmospheric pressure, air temperature, relative humidity, and wind velocity, should be measured at the appropriate time intervals.

The above indirect ET estimates contain many different source of errors, including random errors in meteorological data and propagated errors in ET estimates from the erroneous input.

Knowing these errors as well as sensitivities of ET model will help to understand the structure of ET process as well as to improve a monitoring network for the meteorological data in the region. Sensitivity and error analyses have been extensively studied for rainfall-runoff models [*Main and Brown*, 1978; *Sorooshian and Arfi*, 1982; *Troutman*, 1982, 1985; and others], but those for ET models are rarely found. Thus, the purpose of this study is to define the sensitivity and uncertainty in the ET estimates.

Assuming that the P-B model is a nonlinear regression model where dependent variables are the above mentioned five meteorological variables, the sensitivity and uncertainty analyses were performed with erroneous independent variables measured from an experimental weather station was installed in a south Florida marsh area. In particular, the sensitivity analyses were done with and without the correlation effect of the independent variables. The sensitivity with the correlation effect was evaluated by the conditional probability density function after assuming that two independent variables are a bivariate normal distribution. This result showed that the conditional probability was increased when the correlation effects are accounted for, implying that this case is more suitable for actual meteorological conditions. After defining the measurement errors of the meteorological variables from the multiple measurements, the errors propagated to the ET estimates by the P-B model were computed. The results revealed that both the random measurement error and the propagated errors for net radiation were predominant.

2. Penman-Brutsaert Model

The P-B model, which is based on the classic Penman approach, incorporates the atmospheric drying power with the Monin-Obukhov similarity theory for atmospheric stability. In addition, *Brutsaert* [1979,1982] has developed a theoretical formulation of the scalar roughness lengths for different surface types on the basis of a local Reynolds number for air flow. Thus, one of advantages of using the P-B model is that the scalar roughnesses for different surface conditions can appropriately be accounted in the model so as to eliminate the requirement to parameterize the canopy condition which introduces considerable uncertainty.

The general form of the Penman-Brutsaert combination equation [*Brutsaert and Stricker*, 1979; *Brutsaert*, 1982] is

$$E_p = \frac{\Delta}{\Delta+\gamma}(Rn-G) + \frac{\gamma}{\Delta+\gamma}E_a \tag{1}$$

where E_p is the latent heat flux, Δ is the slope of the saturation vapor pressure-temperature curve, γ is the psychrometric constant, Rn is the net radiation, G is the soil heat flux, and E_a is the drying power of the air. Since the saturation vapor pressure is a function of the ambient air temperature T, Δ is a function of T, and γ is a function of both T and the atmospheric pressure p. Based on the Monin-Obukhov similarity, *Brutsaert* [1982] suggested the drying power of the air as

$$E_a = \kappa u_* \rho (q_a^* - q_a) \left(\ln\left[\frac{z-d_{ov}}{z_{0v}}\right] - \psi_v\left[\frac{z-d_{0v}}{L}\right]\right)^{-1} \tag{2}$$

where κ is von Karman's constant, u_* is the friction velocity of air at the surface, ρ is the density of air, d_{ov} is the displacement height for water vapor, z is the height of measurement of meteorological variables above the surface, z_{ov} is the vapor roughness height, q_a and q_a^* are the specific humidity and the saturation specific humidity at the ambient air temperature, respectively, ψ_v is the stability correction function for vapor, and L is the Obukhov length defined by

$$L = \frac{-u_*^3}{\kappa g[(H+0.61 c_p T E_p)/(\rho c_p T)]} \tag{3}$$

where g is the acceleration of gravity, c_p is the specific heat at constant pressure, and H is the specific flux of sensible heat, which can be expressed from the surface energy budget:

$$H = R_n - G - E_p. \tag{4}$$

The friction velocity is obtained from the mean horizontal wind speed V described in the context of Monin-Obukhov similarity:

$$u_* = \kappa V \left[\ln\left(\frac{z-d_0}{z_0}\right) - \psi_m\left(\frac{z-d_0}{L}\right) \right]^{-1} \tag{5}$$

where d_0 is the momentum displacement height, z_0 is the surface roughness height for momentum, and ψ_m is the stability correction function for momentum.

The atmospheric stability classification is based on L, where $L < 0$, $L > 0$, and $|L| \geq 100$ are the limits used to signify unstable, stable, and neutral conditions, respectively. Then, the stability correction functions for both air and momentum have been obtained from the Businger-Dyer formulation [*Brutsaert*, 1982]. For unstable conditions, these functions are:

$$\psi_m = \ln\left[\frac{(1+x_m)^2(1+x_m^2)}{(1+x_0)^2(1+x_0^2)}\right] - 2\arctan(x_m) + 2\arctan(x_0) \tag{6}$$

$$\psi_v = 2\ln\left(\frac{1+x_v^2}{2}\right) \tag{7}$$

where $x_m=x_v=(1-16y_m)^{1/4}$ with $y_m=(z-d_0)/L$, and $x_0=(1-16z_0/L)^{1/4}$. For stable conditions, both ψ_v and ψ_m are given by

$$\psi_v = \psi_m = 5(z_0/L - y_m) \qquad 0 < y_m \leq 1 \tag{8}$$

$$\psi_v = \psi_m = -5\ln\left(\frac{z-d_0}{z_0}\right) \qquad 1 < y_m \leq 10. \tag{9}$$

Brutsaert [1982] suggested that the roughness heights z_{ov} and z_{oh} depend on the surface condition with the roughness Reynolds number $z_{0+}(=u_*z_0/\upsilon)$, where υ is the kinematic viscosity of air. That is, for smooth surfaces ($z_{0+}<0.13$), the scalar roughness values are:

$$z_{0v} = 0.624\frac{\upsilon}{u_*} \tag{10}$$

$$z_{0h} = 0.395\frac{\upsilon}{u_*} \tag{11}$$

or for the bluff-rough surface condition ($z_{o+}>2$), they both z_{ov} and z_{oh} are:

$$z_{0v} = 7.4z_0\exp(-2.25z_{0+}^{1/4}) \tag{12}$$

$$z_{0h} = 7.4z_0\exp(-2.46z_{0+}^{1/4}). \tag{13}$$

For the transitional flow regime between smooth and bluff-rough conditions ($0.13<z_{o+}<2$), no theoretical expressions for both z_{ov} and z_{oh} are available, but Merlivat's criterion of $z_{o+}=1$ is known to be acceptable as a transition point from smooth to bluff-rough surfaces [*Merlivat, 1978; Brutsaert*, 1982].

Since the equations (1) through (5) are not explicit, they should be solved simultaneously. *Katul and Parlange* [1992] found that a simple iterative procedure gives a good convergent solution for E_p. Having a onset latent heat flux rate E_p, the potential ET rate (L/T) is computed by ($ET=E_p/\lambda$), where λ is the latent heat of vaporization that is a function of the air temperature T. The specific humidity q_a in (2) is a function of the actual vapor pressure e_d that is a function of the relative humidity RH. The direct measurement of soil heat flux G requires a heat plate installation and calibration of instruments needing great care [*Brutsaert*, 1982]. Since G is usually not sensitive to the estimated ET rate (discussed later), the following simple empirical relationship was adopted to estimate G:

$$G = c_r Rn \tag{14}$$

where parameter c_r needs to be calibrated based on the actual ET. If the height of vegetation h_c is available, both the displacement height d_{ov} and the surface roughness height for momentum z_{om} can not defined theoretically, but if vegetation is dominant at the surface, the following empirical relationships are commonly used in practice:

$$d_{0v} = d_0 = c_d h_c \tag{15}$$

$$z_{0m} = c_z h_c \tag{16}$$

where h_c is the height of vegetation, and c_d and c_z are the parameters to be determined.

To summarize the above procedure, a set of input vector x to the P-B model is x={Rn, p, T, RH, and V}, and a parameter set β that needs to be defined in advance is β={h_c, z, c_z, c_d, and c_r}. In addition, both z and h_c should be known in order to estimate ET rate by the P-B model.

3. Sensitivity Analysis for Correlated Independent variables

Sensitivity analysis investigates the changes in the optimal solution with the changes in the optimal system input components y, such as the input variables x or the parameters β. The results of the sensitivity analysis can be used to track and account for errors in a model simulation and to characterize the resulting ranges of uncertainties. The sensitivity and uncertainties in ET predictions may be analyzed by treating the P-B model as a non-linear regression model. That is, the P-B model utilizes the independent variables x and the fixed parameters β to compute the ET; the optimal ET rate predicted by the P-B model can now be simply denoted by $ET^0=f(y^0)=f(x^0,\beta^0)$, where the superscript "0" denotes the optimal system input in contrast to the erroneous one. Letting y_i^0 be the i-th input component among vector y^0, the sensitivity s_i by changing y_i^0 is expressed by the gradient term [*Singh*, 1988]:

$$s_i = \frac{\Delta f(.)}{\Delta y_i} = \frac{f(y_i,\, y_j^0,\, j\neq i) - f(y_0)}{y_i - y_i^0} \tag{17}$$

where $y_i=y_i^0+\Delta y_i=y_i^0(1+\delta/100)$. The common values of δ may be (±5%, ±10%, ±15%,.....). Sensitivity s_i can also be defined by the partial derivative term $\partial f(.)/\partial y_i$. Both partial derivative and gradient sensitivities are identical only when the relationship of dependent and independent variables are linear. More often than not, the above gradient sensitivity is more preferable than that of partial derivative, since the variations of both dependent and independent variables are more adequately represented by the gradient sensitivity.

If independent variables are correlated (which is true for some independent variables in the P-B model), the correlation effect may be included in the sensitivity analysis as follows: Let us consider that, for any given model, the sensitivity by changing the i-th optimal independent variable x_i^0 is of concern and that the j-th independent variable x_j is significantly correlated with x_i. For a given x_i, x_j may be more than one. It is obvious that there exist some independent variable(s) whose correlations are not very significant. Then, the sensitivity s_i with respect to changing x_i^0 with the optimal parameters set β^0 is

$$s_i = \frac{\Delta f(.)}{\Delta x_i} = \frac{f(x_i,\ E[X_j^0|X_i=x_i],\ x_k^0,\ j\neq i, k\neq i, k\neq j,\ \beta^0) - f(x^0,\beta^0)}{x_i - x_i^0} \tag{18}$$

where $x_i=x_i^0+\Delta x_i=x_i^0(1+\delta/100)$, and uppercase ($X_j$) is the random variable while lowercase (x_j) is any specified value. It is also possible to extend the above formulation to the multi-variate case rather than the bi-variate case, that is, multi-variate x_i's with multiple x_j, which is not included in this paper. To obtain the conditional expectation in (18), it is assumed that a simple and reasonable model which represents adequately the relationship between x_j and x_i is

$$x_j = \alpha_0 + \alpha_1 x_i + \epsilon \tag{19}$$

where α_0 and α_1 are the regression parameters, and ε is the regression error which is independent of x_i having $\varepsilon \sim N(0,\sigma_\varepsilon^2)$. Further, it is assumed that both x_j and x_i are a bivariate normal distribution having a correlation coefficient ρ. Then, the conditional expectation and variance of x_j given x_i are given [*Mood et al.*, 1974, Chapter 5] by

$$E[X_j^0|X_i=x_i] = \alpha_0 + \alpha_1 x_i = \mu_{X_j} + \frac{\rho\sigma_{X_j}}{\sigma_{X_i}}(x_i-\mu_{X_i}) \tag{20}$$

$$Var(X_j^0|X_i=x_i) = \sigma_\epsilon^2 = (1-\rho^2)\sigma_{X_j}^2 \tag{21}$$

where $\mu_{(.)}$ and $\sigma_{(.)}^2$ are the sample mean and variances of the given independent variable. If the conditional expressions on x_j given x_i are already known, the conditional expressions of x_i given x_j are simply given [*Troutman*, 1982] by

$$E[X_i^0|X_j=x_j] = \frac{\mu_{X_i}\sigma_\epsilon^2 - \alpha_0\alpha_1\sigma_{X_i}^2 + \alpha_1\sigma_{X_i}^2\, x_j}{\sigma_\epsilon^2 + \alpha_1^2\sigma_{X_i}^2} \tag{22}$$

$$Var(X_i^0|X_j=x_j) = \frac{\sigma_\epsilon^2\sigma_{X_i}^2}{\sigma_\epsilon^2 + \alpha_1^2\sigma_{X_i}^2}. \tag{23}$$

The sensitivity result obtained by equation (18) can be evaluated by computing the conditional probability density function (pdf) of x_i given x_j, or vice verse, and by comparing it with the uncorrelated one. If a set of two random variables (X_i, X_j) is a bivariate normal distribution and its correlation coefficient ρ is significant, the conditional distribution of X_i given $X_j=x_j$ is normal with mean $\mu_{X_i}+(\rho\sigma_{X_i}/\sigma_{X_j})(x_j-\mu_{X_j})$ and variance $\sigma_{X_i}^2(1-\rho^2)$, and the conditional pdf of x_i given by x_j is obtained from its joint and marginal pdfs [*Mood et al.*, 1974] as

$$f_{(X_i|X_j)}(x_i|x_j) = \frac{f(x_i,x_j)}{f_{X_j}(x_j)} = \frac{1}{\sqrt{2\pi}\sigma_{X_i}\eta}\exp\left\{-\frac{1}{2\sigma_{X_i}^2\eta^2}\left[x_i-\mu_{X_i}-\frac{\rho\sigma_{X_i}}{\sigma_{X_j}}(x_j-\mu_{X_j})\right]^2\right\} \tag{24}$$

where $\eta^2=(1-\rho^2)$. In the case of equation (17) which is the sensitivity without the correlation effect, x_j^0 is simply taken by its mean μ_{X_j}, then the above conditional pdf becomes

$$f_{(X_i|X_j)}(x_i|X_j=x_j^0) = \frac{1}{\sqrt{2\pi}\sigma_{X_i}\eta}\exp\left\{-\frac{1}{2\sigma_{X_i}^2\eta^2}(x_i-\mu_{X_i})^2\right\}. \tag{25}$$

In the case of equation (18) which is the sensitivity with the correlation effect, the regression curve of x_j given by x_i is the straight line obtained by equation (20) [*Mood et al.*, 1974]. Then, plugging equation (20) into equation (24) results in the conditional pdf of

$$f_{(X_i|X_j)}(x_i\ |X_j=E[X_j^0|X_i=x_i]) = \frac{1}{\sqrt{2\pi}\sigma_{X_i}\eta}\exp\left\{-\frac{1}{2\sigma_{X_i}^2}(x_i-\mu_{X_i})^2\right\} \tag{26}$$

which is identical to the marginal pdf of x_i divided by η, that is, $f_{Xi}(x_i)/\eta$. Equation (22) is the conditional pdf parallel to the x_i axis while equation (24) is the conditional pdf parallel to the regression line given by equation (21). Even though sensitivities obtained by equations (17) and (18) have different meanings, the higher conditional pdf ensures more reliability of the sensitivity analysis. If ρ is significant, the conditional pdf obtained by equation (26) is always greater than that by equation (25) for $|x_i|>0$, implying that the sensitivity analysis obtained by equation (18) represents more for the actual sample than that by equation (17).

4. Error Analysis

Predicted ET may contain many different sources of errors, including model error, temporal and spatial scaling error, error from biased parameters, and propagated error due to erroneous independent variables. The general assumptions in error analysis are: (1) the errors are statistically independent of the predictions $f_i(x,\beta)$ for different time i, and are identically distributed; (2) the errors are statistically independent from each other; and (3) the errors are normally distributed, with a zero mean and a finite variance [*Troutman*, 1985].

The convention is to let random variables be uppercase (Y,X) and particular observations lowercase (y,x). Given a set of error-free independent variable x^*, in contrast to the erroneous independent variable x, there is a range of actual values of ET that could be conceivably be associated with x^*; this range is characterized by the probability distribution of ET conditioned on $X^*=x^*$, or $(ET|X^*=x^*)$. Letting the actual ET as $(ET|X^*=x^*)$ and the model-predicted ET as $f(x^*,\beta^*)$, there will always be a nonzero difference between $(ET|X^*=x^*)$ and $f(x^*,\beta^*)$. Let us define the nonzero differences as the prediction errors e, that is, $e=(ET|X^*=x^*)-f(x^*,\beta^*)$. The e series in time is normally distributed with a mean of zero and a finite error variance of σ_e^{*2}. Using the above definitions, it is possible that the following two conditional expectations are applicable in describing the error process [*Troutman*, 1985]:

$$E[ET|X^*=x^*] = f(x^*,\beta^*) \tag{27}$$

$$Var(ET|X^*=x^*) = E[\{(ET|X^*=x^*)-f(x^*,\beta^*)\}^2] = \sigma_e^{*2} \tag{28}$$

Even though the above expressions are based on the unbiased independent variables x^*, the available data in the field are usually erroneous data x due to the small number of observation, or only a single measurement in most practical cases.

Since errors are normally distributed with a finite variance, the variance of error has been frequently used as a quantity of error. Instead, this paper uses the percentage probabilistic error, which is more generally accepted in the practical metrology to define the error quantity [Rabinovich, 1993]. For a given erroneous independent variable X=x, the total ET prediction error variance $\sigma_p^2(x,\beta)$ can be expressed [*Troutman*, 1982] by

$$\sigma_p^2(x,\beta)=Var(ET|X=x) = E[(ET-f(X,\beta))^2|X=x] = \sigma_e^{*2} + \upsilon^2(x) + \gamma^2(x,\beta) \tag{29}$$

where the first term of the last expression is the variance of model predicted error, and

$$\upsilon^2(x) = Var[f(X^*,\beta^*)|\ X=x] \tag{30}$$

is the propagated error variance from the erroneous independent variables, and

$$\gamma(x,\beta) = E[ET|X=x]-f(x,\beta) \tag{31}$$

is the bias in prediction due to the erroneous parameters. The third term in equation (29) is the expected square bias which purely depends on the model parameters. If the error-free independent variables ($x=x^*$) are used in prediction, $\upsilon^2(x)$ will be zero, or if the correct parameters ($\beta=\beta^*$) are used, the bias is zero and the third term in equation (29) will vanish.

5. Measurement of meteorological data

An experimental weather station was installed at the Everglades Nutrient Removal project site in south Florida ($26^0 38'$ N, $80^0 25'$ W). The project area is primarily an open, treeless, and undeveloped swamp covered with the tall dense cattails. The cattail (Typha domingensis) is about 1.5 meters in height. The elevation of the site is about 4 meters above the National Geodetic Vertical Datum. The area is transacted by small lateral canals and dirt roads which may impact on the aerodynamic process at the interfacial sublayer, which is the immediate vicinity of the surface where the turbulence of air flow is strongly affected by surface roughness elements.

Three 10-meter wind towers and one 2.6-meter wind tower (Figure 1) were installed, from which wind velocity (V) and direction have been measured at 15-minute intervals. Each 10-

meter wind tower has three identical sets of weather sensors installed 2 meters above land surface (z=2m) to measure the meteorological data, including Rn, p, T, RH. Accordingly, each meteorological variable has 9 multiple measurements (WXij, i=1,2,3, j=1,2,3), while wind speed has three 10-meter measurements (WX11, WX21, and WX31) and one 2.6-meter measurement (WX12). The distances between the three wind towers are less than 100 meters. The statistics (which are discussed later) show that no significant differences were found between the meteorological data from three wind towers, indicating that weather conditions at the three wind tower are nearly identical. Also, *Famiglietti and Wood* [1995] showed that the representative element area, which is the critical scale at which the implicit continuum assumption can be used, for diurnal areally averaged ET reaches a maximum of 1-2 km^2. Therefore, multiple measurements of each meteorological data were treated as a point data and differences between them were considered as random errors. Since the number of multiple measurements is small, the unbiased estimation method was used to compute the statistics of the meteorological data.

The weather station has a non-weighable submerged lysimeter. The lysimeter has a diameter of 3.5-meters and is planted in cattail to mimic the surrounding cattail field. The water budget components, including stage, inflow, outflow, and rainfall within the lysimeter, have been recorded at 15-minute intervals to compute the lysimeter ET which may be used to calibrate and verify any conceived ET models. Therefore, the lysimeter ET rates are structurally independent from the ET rates estimated by the meteorological data. However, the 15-minute lysimeter ET rates include considerable error due to the low accuracy of the stage measurement: the measurement accuracy of the lysimeter stage is about 0.38 millimeters whereas an average estimated daytime ET rate is about 0.15 millimeters per 15 minutes. The lysimeter errors are reduced when the 15-minute ET rates are aggregated to a longer time step such as daily (will be discussed later).

6. Results and Discussion

Sensitivity without the correlation effect

Both sensitivity and error analyses were based on one-year (10/14/1992-10/13/1993) of meteorological data from the above experimental weather stations. As a preliminary analysis, the basic statistics of a set of unbiased independent variables x^* were computed. To explain this procedure, let us define a three-dimensional array of erroneous independent variables whose element is $\{x_{ijk}\}$, where

- $i=1,\ldots,I$, with I as the number of independent variables (I=5),
- $j=1,\ldots,J$, with J as the number of time steps (J=35040 for overall year, or J=17521 for day-time only), and
- $k=1,\ldots,K$, with K as the number of multiple measurements (K=9 for $i=1,\ldots,4$, or K=3 for i=5 which is wind velocity V).

After assuming that the multiple measurements of each variable at each time are normally distributed, the unbiased independent variables vector x_{ij}^* were computed by taking the arithmetic mean, that is $x_{ij}^* = (x_{ij1}+\ldots+x_{ijK})/K$. Then, the basic statistics of each independent variable were computed along the time axis, i.e. $j=1,\ldots,J$ (Table 1). The daytime was assumed when the 15-minute net radiation readings were positive, that is, $Rn_{ij}^*>0$. The portion of daytime is about 50 percent of the period of record. Particularly, the mean squared error, MSE_{x_i}, in Table 1 is computed by

$$MSE_{x_i} = \frac{1}{JK}\sum_{j=1}^{J}\sum_{k=1}^{K}(x_{ijk}-x_{ij}^{*})^2. \qquad (32)$$

The interesting thing to note in this table is that the skewness values of the day-time Rn and V are reduced significantly, compared to those of the overall year. These decreased skewness values are good for the sample's normality assumption in both sensitivity and error analyses. As a matter of fact, the night-time ET rates are not significant compared to the daily ET rate; the average night-time net radiation (which is negative) during the period of record is only 7.4 percent of the day-time net radiation. Therefore, most analyses here are those of the daytime data unless specified.

For the sensitivity analyses, the optimal independent variables x_i^0 in equations (17) and (18) were assumed the averaged error-free daytime values x_i^* those of which are listed on the seventh row from the bottom in the Table 1. Using the P-B model with h_c=1.5m, z=2.0m, c_z=0.123 , c_d=2/3 , and c_r=0.0817, the estimated ET rate is 7.413 mm per day, where c_r was calibrated using the actual daily lysimeter ET rates and both c_z and c_d were taken from *Jensen et al.* [1990]. The range of changing independent variable Δx_i in equations (17) and (18) was from $(-1.5\sigma_{X_i})$ to $(+1.5\sigma_{X_i})$. For plotting purposes, the x_i's were scaled by the standardization process to make rational comparisons of the changing estimated ET rate by changing the each independent variable based on its standard deviation. That is, with μ_{X_i} and σ_{X_i} the mean and standard deviation of i-th independent variable, respectively, the standardized variable x_i' is

$$x_i' = \frac{x_i-\mu_{X_i}}{\sigma_{X_i}} \qquad (33)$$

which has a mean of zero and a variance of one. Figure 2 shows the resulting sensitivity curves of five independent variables. The results showed that, among the five independent variables, Rn was the most sensitive to the predicted ET rate, and that RH was the second most sensitive but its magnitude was much less than that of Rn, and that both T and V were nonlinearly proportional to the ET rate but their sensitivities were relatively insignificant.

Figure 3 shows the sensitivity curves of five major parameters defined in the P-B model, where δ which is the changing rate from the optimal parameters was from -20% to 20%. The sensitivity curves for h_c, z, c_z, and c_d were nonlinear due to the fact that those parameters reflect the aerodynamic process in the ET model. In general, the above five parameters were usually much less sensitive to the ET prediction compared to those of independent variables. The present analyses were those of a small spatial scale (virtually a point estimate). However the sensitivities of the parameters are expected to increase in proportional to the increasing spatial scale.

Sensitivity with the correlation effect

Table 2 shows correlation coefficients between the five error-free independent variables, as well as the latent heat flux E_p and the sensible heat flux H estimated by the P-B model. This matrix revealed that the correlation between Rn and RH in the humid region was higher than those of any other combinations of independent variables. The relative humidity RH, which is the ratio (expressed as a percentage) of the actual to the saturation vapor pressure in the air, is an inverse function of the air temperature which is again a function of Rn. Unlike the actual vapor content in the air, RH in the area is at a minimum in the afternoon and at a maximum in the early morning. The matrix also showed that both E_p and H were highly correlated with Rn, implying that ET rates can be predicted simply by only Rn with considerable accuracy for daily or larger time steps. The correlation coefficients of the overall data were slightly higher than those of day-time data, but those of the overall data are biased since the night-time Rn and T are clustered to their low values.

Taking into account for the high correlation between Rn and RH, the sensitivity analyses with correlation effect were done for those two variables as follows: First, the regression equations and its error variance were estimated by equations (20) through (23) as:

$E[RH^0|Rn] = 87.319\text{-}48.568Rn$ with $\sigma_\epsilon^2=153.25$, and

$E[Rn^0|RH] = 0.898\text{-}0.00851RH$ with $\sigma_\epsilon^2=0.02686$,

where units of RH and Rn are percent and kw/m^2, respectively. Then, the sensitivities were computed by equation (18) while changing each independent variable ranging from -1.5σ to 1.5σ, whose results are plotted in Figure 4. With the correlation effect between Rn and RH, the sensitivity for RH was dramatically increased, while that for Rn was moderately increased. Nevertheless, the sensitivity for Rn was still higher than that for RH.

In order to evaluate the sensitivity analyses with and without the correlation effect, the conditional pdf's were computed by equations (25) and (26), and plotted in Figure 5. Under the bivariate normal distribution assumption of Rn and RH, the conditional pdf's with the correlation effect were increased compared to those without the correlation effect. The result implies that, even though each case has its own meaning, the sensitivity with the correlation effect is more resemble to the actual sample in probabilistic sense than that without the correlation effect. Also, it should be noted that, the conditional pdf of Rn given RH was higher than that of RH given Rn, due to the fact that the variance of the sample RH in the humid region was relatively less than that of Rn (refer to Table 1).

Error analysis

The total prediction error variance $\sigma_p^2(x,\beta)$ of any given model can be defined by equation (29) which has three components; the model error, propagation error, and error due to the biased parameters. Having all required meteorological data, only c_r for G is needed to calibrate for the use of the P-B model. Noticing that the sensitivity of c_r is relatively low; it does not induce any significant bias on the estimated ET rate. Thus, it is assumed that $\gamma^2(x,\beta)$ term (29) is zero and the first two terms are defined here. To investigate the independence in the ET estimation errors in time, the auto-correlation coefficients of error series were computed, which are 0.37 and -0.21 for the first and second time lags, respectively. Those low auto-correlations indicate that the ET estimation errors were independent in time. Further sophisticated test for correlated errors are also available [*Sorrooshian and Dracup*, 1980; *Alley*, 1984], but are not considered in here since the auto-correlations of prediction errors were insignificant.

As mentioned before, the 15-minute lysimeter ET rates contain considerable errors which are usually more than 100%, and are greater during the day-time. Thus, the first components in equation (29), the model error variances, were computed on a daily basis as follows: the daily sum ET series were computed from the 15-minute water budget components in the lysimeter which are available only after February 11, 1993 (245 days among the period of record of meteorological data). Also, the ET rates during the same periods were estimated by the P-B model in 15-minute interval and aggregated to get the daily estimated ET series. From both estimated and measured daily ET series, the variance of model errors σ_e^{*2}, as well as those of rainy and no-rain days were computed (Table 4). Noticing that the ET estimation error by the P-B model is usually small for even small time intervals [*Stricker and Brutasert*, 1978; *Katul* *and Parlange*, 1992; and *Parlange and Katul*, 1992], the model errors in Table 4 are the errors caused by the inaccurate lysimeter measurements. From σ_e^{*2}'s with and without rainfalls, it can be concluded that the considerable error in the daily ET estimations originated from the uncertainty of rainfall: the uncertainty in rainfall causes about a 28% increase of the model error variance. However, the majority source of error in σ_e^{*2} was the lysimeter measurement error.

The second component in equation (29), the propagated error variances $\upsilon^2(x)$, were computed by equation (30) for each set of erroneous 15-minute independent variable x_{ik} ($i=1,...,5$, $k=1,...9$ or 3 for wind speed), and the results are listed in Table 5. This table also includes the average $\upsilon^2(x)$'s for both 15-minute and daily intervals. The results indicate that the propagated errors $\upsilon^2(x)$'s from variable to variable changes significantly, that the erroneous Rn resulted in the most severe error in the ET estimates while the effect of p, T, RH, and V were relatively low. The erroneous RH induced the second most significant error in ET estimates, and p was the most insignificant. The overall results of the propagated errors were very similar to that of the sensitivity analysis even though they are different in nature. The aggregated daily propagated errors from those of 15-minute were reduced significantly (last row in Table 5) compared to those of 15-minute, due to the fluctuation of 15-minute errors along the zero mean in time. In other words, the propagated errors of 15-minute data were unbiased along the mean. Also, the computed $\upsilon^2(Rn,p,T,RH,V)$, which is the propagated error variance of 15-minute ET rates resulting from the combined effect of all five independent variables (Rn, p, T, RH, and V), was 1.038 and that of daily interval was 0.097.

As a summary, Table 6 lists the percentage probabilistic errors ξ to the expected independent variable X_i as

$$\xi = \frac{100 t_q S_e}{\mu_{X_t}} \tag{34}$$

where S_e the unbiased sample standard deviation of any type of errors, and t_q is the q percent point of the Student's t distribution depending on the confidence probability level α and the degree of freedom ν=n-1 with n is the number of samples. If the above quantities are for ET rates, μ should be an average ET rate of 7.413 mm/day. The above probabilistic error implies that the most errors e's in time with probability level α falls within the limits $\pm\xi$ [*Rabinovich*, 1993]. The Table 6 showed that both measurement and propagated errors for Rn were most significant, while those of p were minimal. However, it should be noted that the percentage probabilistic error of υ^2 (V) were greater than that of υ^2 (RH) due to the significant measurement errors on V, whereas the sensitivity results of them were opposite.

7. Conclusions

1. For purposes of the sensitivity analysis in the Penman-Brutsaert (P-B) evapotranspiration model, a theoretical framework of sensitivity analysis method for correlated independent variables was formulated, along with the verification method by the conditional probability density function (pdf) after assuming that any sets of the correlated two independent variables are bivariate normally distributed. Whenever the correlations of any combinations of two independent variables are significant, the conditional probability with correlation effect increases, compare to that without correlation effect, with proportion to the increasing correlation coefficient between two variables. This method has the potential to be applied to the sensitivity analysis of any given model with significant correlations presented in the independent variables.

2. The sensitivity analyses with 15-minute meteorological data indicate that the net

radiation Rn was the most sensitive to the ET estimate, that relative humidity RH was the second most sensitive in the humid region and was inversely proportional to the ET rate, and that temperature T and wind velocity V were less sensitive and were nonlinearly behavior to the estimated ET rate. The statistical analyses show that Rn was highly correlated to the latent heat flux E_p (ρ=0.99). These results may change depending on the temporal and spatial scales which should be further studied in terms of spatial variability of the ET rate.

3. The measured data show that the correlation coefficient between Rn and RH was significant (ρ=-0.74) in the swamp region where the potential ET rates are identical to the actual ET rates. The sensitivity analysis with correlation effect shows that the sensitivity of estimated Et rate by changing RH was significantly increased, whereas that of Rn was mildly increased but the sensitivity for Rn was still higher than for RH. The conditional pdf's of both (Rn|RH) and (RH|Rn) were increased, but the former case was more significant than the later owing to the small variation of RH. This result indicates that, even though both sensitivities with and without the correlation effect have their own meaning, the sensitivity with the correlation effect gives more assurance than that without the correlation effect.

4. The random measurement errors for five independent variables (Rn, p, T, RH, and V) to the P-B model were defined from the multiple samples at the experimental weather station in the south Florida. Those random errors of meteorological data were well above of the instrumental errors. The propagated error variance $\upsilon^2(x)$ from the erroneous independent variables were obtained, which are not available in actual measurement cases where single measurement at a time is in common. Both measurement and propagated errors will be valuable for the interpretation of the estimated ET rates as well as the design of further meteorological

monitoring networks.

5. The percentage probabilistic errors to the expected variable based on the normality assumption were presented (Table 5) for two different probability levels (α of 0.75 and 0.9). The result shows that the propagated error in ET rate estimated from the erroneous net radiation was most significant; the expected propagation error in the estimated ET rate due to the erroneous Rn is about 11 percent when α is 0.75. The percentage probabilistic error for V is higher than that for RH making it second most significant variable due to the large measurement error on V, whereas the sensitivity results of both RH and V were opposite.

Acknowledgments.

The author acknowledges Duane Dunn, Ken Perla, and Eric Zhang, all in Data Management Division, South Florida Water Management District, for providing me with the weather station data.

References

Alley, W. M., On the treatment of evapotranspiration, soil moisture accounting, and aquifer recharge in monthly water balance models, Water Resources Research, 20(8), 1137-1149, 1984.

Brutsaert, W., Heat and mass transfer to and from surfaces with dense vegetation or similar permeable roughness, *Boundary Layer Meteorology*, 16, 365-388, 1979.

Brutsaert, W., *Evaporation Into Atmosphere: Theory, History and Applications*, D. Reidel, Hingham, Mass., 1982.

Brutsaert, W. and H. Stricker, An advection-aridity approach to estimate actual regional evapotranspiration, Water Resources Research, 15. 443-450, 1979.

Famiglietti, J. S. and E. F. Wood, Effect of spatial variability and scale on areally averaged evapotranspiration, *Water Resources Research*, 31(3), 699-712, 1995.

Katul, G. G. and M. B. Parlange, A penman-Brutsaert model for wet surface evaporation, *Water Resources Research*, 28(1), 121-126, 1992.

Jensen M. E., R. D. Burman, and R. G. Allen (Edited), *Evapotranspiration and Irrigation Water Requirements*, ASCE Manuals and Report on Engineering Practice No. 70, ASCE, N.Y., 1990.

Jones, J.W., L.H. Allen, S.F. Shih, J.S.Rogers, L.C. Hammond, A.G. Smahstrala, and J.D. Martsolf, Estimated and measured evapotranspiration for Florida climate, crops, and soils, *Bulletin 840* (technical), Agricultural Experimental Stations, Inst. of Food and Agr. Sci., Univ. of Florida, Gainesville, FL, 1984.

Main, R.G., and B.M. Brown, Sensitivity of optimized parameters in watershed models, *Water*

Resources Research, 14(2), 299-303, 1978.

Merlivat, L., The dependence of bulk evaporation coefficients on air-water interface conditions as determined by the isotopic method, J. Geophysical research (Oceans and Atmos.) 83(C6), 2977-2980.

Mood, A. M., F. A. Graybill, and D. C. Boes, *Introduction to the Theory of Statistics*, 3rd Edition,McGraw-Hill Book Co., N. Y., 1974.

Parlange, M. B. and G. G. Katul, An advection-aridity evaporation model, *Water Resources Research*, 28(1), 127-132, 1992.

Rabinovich, S., *Measurement Errors: Theory and Practice*, American Institute of Physics, N. Y., 1993.

Singh, V. P., *Hydrologic System: Volume I. Rainfall-runoff Modeling*, Prentice Hall, Englewood Cliffs, N.J., 1988.

Sorooshian, S., and F. Arfi, Response surface parameter sensitivity analysis methods for postcalibration studies, *Water Resources Research*, 18(5), 1531-1538, 1982.

Sorooshian, S., and J. Dracup, Stochastic parameter estimation procedures for hydrologic rainfall-runoff models: Correlation and heteroscedastic error case, *Water Resources Research*, 16(2), 430-442, 1980.

Stannard, D. I., Comparison of Penman-Montieth, Shuttleworth-Wallace, and modified Priestly-Taylor evapotranspiration models for wildland vegetation in semiarid rangeland, *Water Resources Research*, 29(5) 1379-1392, 1993.

Stricker, H., and W. Brutsaert, Actual evapotranspiration over a summer period in the "Hupsel Catchment", Journal of Hydrology, 39, 139-157, 1978.

Troutman, B. M., An analysis of input errors in precipitation-runoff models using regression with errors in the independent variables, *Water Resources Research*, 18(4), 947-964, 1982.

Troutman, B. M., Errors and parameter estimation in precipitation-runoff modeling, 1. Theory, *Water Resources Research*, 21(8), 1195-1213, 1985.

Winter, T. C., D. O. Rosenberry, and A. M. Sturrock, Evaluation of 11 equations for determining evaporation for a small lake in the north central United States, *Water Resources Research*, 31(4), 983-993, 1995.

Table 1. Statistics of the Error-free Independent Variables Used in Evapotranspiration Estimation.

Statistics	Rn (kw/m^2)	p (kpa*)	T (°C)	RH (%)	V (m/s)
Overall Year (10/14/1992-10/13/1993; 35040 samples)					
Mean	0.122	101.719	22.834	83.495	1.103
Standard deviation	0.208	0.328	5.350	15.446	0.795
Skewness	1.363	-0.467	-0.381	-0.907	1.280
Minimum	-0.071	99.530	1.833	26.861	0.003
Maximum	0.880	103.080	35.717	99.935	8.008
Coeff. of variance	1.702	0.003	0.234	0.185	0.721
Mean square error	0.001	0.223	0.066	4.709	0.009
Daytime only (10/14/1992-10/13/1993; 17521 samples)					
Mean	0.264	101.718	25.509	74.497	1.394
Standard deviation	0.214	0.336	4.752	16.164	0.880
Skewness	0.517	-0.533	-0.549	-0.059	0.872
Minimum	0.000	99.530	2.659	26.861	0.003
Maximum	0.880	103.080	35.717	99.9350	8.008
Coeff. of variance	0.808	0.003	0.186	0.217	0.631
Mean square error	0.002	0.246	0.080	3.404	0. 010

* 1 kpa (kilopascal) = 10 milibar

Table 2. Correlation Coefficient Matrix of Independent Variables, as well as Latent Heat Flux E_p and Sensible Heat Flux H.

	Rn	p	T	RH	V	E_p	H
Rn	1.0	0.032	0.586	-0.740	0.351	0.994	0.859
p	0.073	1.0	-0.135	-0.038	-0.096	0.022	0.068
T	0.514	-0.095	1.0	-0.456	0.120	0.597	0.432
RH	-0.643	-0.065	-0.372	1.0	-0.495	-0.783	-0.415
V	0.178	-0.098	-0.046	-0.342	1.0	0.414	0.003
E_p	0.990	0.054	0.550	-0.700	0.264	1.0	0.797
H	0.802	0.132	0.241	-0.250	-0.224	0.709	1.0

Ref.: The elements in the upper triangular matrix is for the overall year data, whereas those of the lower triangular matrix is for the daytime data

Table 3. Accuracy of Instruments Used at the Experimental Weather Station.

Independent Variable	Instrument Type	Accuracy
Net Radiation, Rn	Q-6 Net Radiometer	0.02 kw/m^2
Atmospheric Pressure, p	PTA-427 Barometric Pressure Transducer	0.133 kpa
Air Temperature, T	HMT 35-C Sensor	0.5 ^{0}C
Relative Humidity, RH	HMT 35-C Sensor	0.5 %
Wind Velocity, V*	W. Tronics 2100 Skyvane	0.45 m/s

* V was measured every 10 seconds and averaged over 15-minute intervals.

Table 4. Measurement Errors in the Daily Lysimeter ET Rate.

Type of Error	Value
Measurement error in the Lysimeter Stage[+]	0.38 mm
Mean Absolute Error	1.411 mm/day
σ_e^{*2} (245 days)	3.172 $(mm/day)^2$
σ_e^{*2} for rainy days (93 days)	3.669 $(mm/day)^2$
σ_e^{*2} for no-rain days (152 days)	2.867 $(mm/day)^2$

+ Stage measurement error is the instantaneous error in each measurement (15-minute intervals).

Table 5. Propagated Error Variances $\upsilon^2(x)$ by the Erroneous input Variables and Their Averages

Station ID	$\upsilon^2(Rn)$ $\times 10^0$	$\upsilon^2(p)$ $\times 10^5$	$\upsilon^2(T)$ $\times 10^3$	$\upsilon^2(RH)$ $\times 10^3$	$\upsilon^2(V)$ $\times 10^3$
wx11	1.035	1.683	1.324	5.074	5.438
wx12	1.308	0.406	1.366	2.736	-
wx13	0.820	0.041	1.080	1.879	-
wx21	0.956	0.120	0.903	1.189	4.966
wx22	0.662	0.154	1.213	3.642	-
wx23	0.846	0.084	1.047	1.404	-
wx31	0.997	0.410	1.317	3.378	4.441
wx32	0.596	0.998	2.722	2.203	-
wx33	1.385	0.065	1.388	4.406	-
Average	0.956	0.440	1.388	2.879	4.948
1-day Average	0.083	0.025	0.100	0.430	0.090

Ref.: The above error were estimated in 15-minute intervals except for the last row, and the unit of estimated ET rate was mm/day.

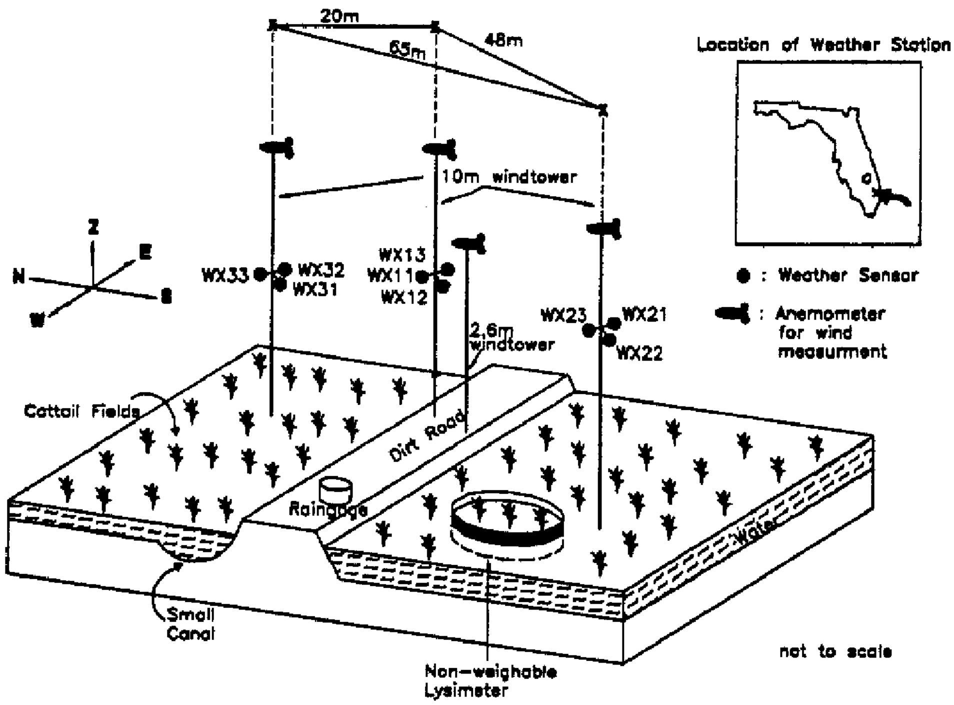
Table 6. Percentage Probabilistic Errors.

unit: %

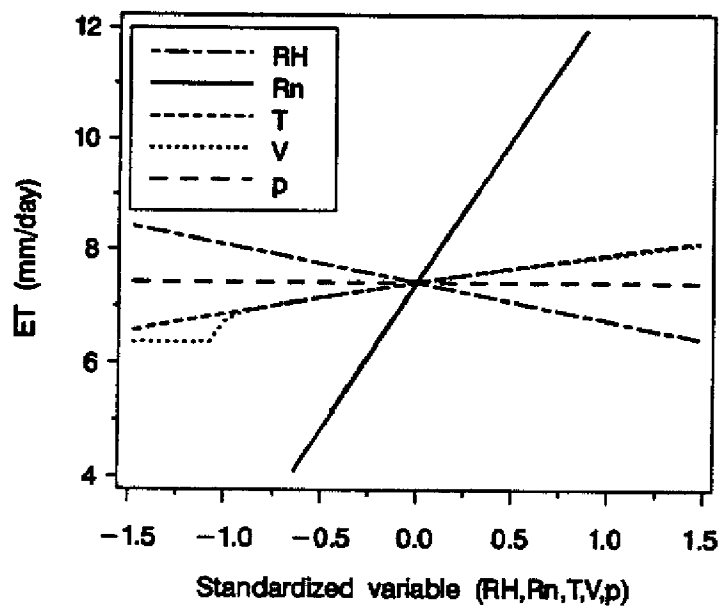
Error Type	Prob.	Rn	p	T	RH	V	(Rn,p,T,RH,V)
1. Random Error in Input	α=0.75	14.65	0.33	0.75	1.67	4.84	-
	α=0.90	21.70	0.63	1.42	3.18	9.20	-
2. Propagated Error, $\upsilon^2(x)$	α=0.75	8.991	0.02	0.34	0.49	0.64	9.33
	α=0.90	16.91	0.04	0.64	0.93	1.22	17.95
3. Model Error, daily σ_e^{*2}	α=0.75	-	-	-	-	-	11.46
	α=0.90	-	-	-	-	-	21.80

Ref.: α or prob. is the confidence level; all errors except the last row are for 15-minute intervals data.

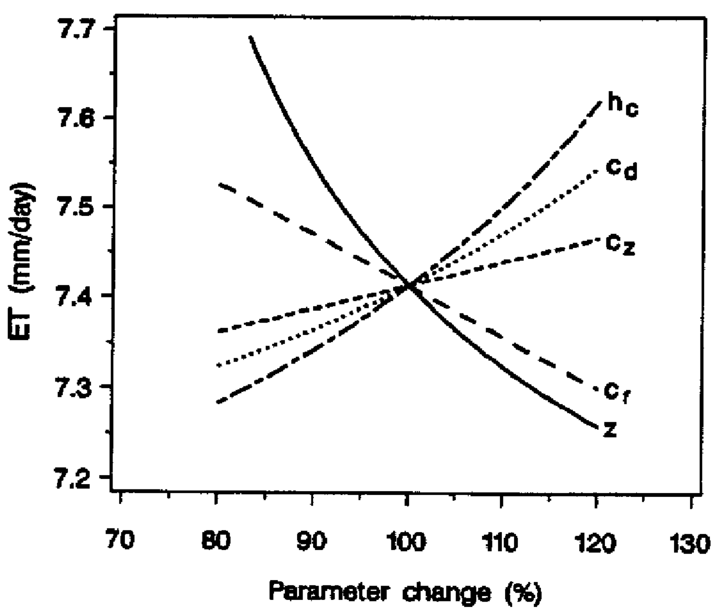
List of Figures

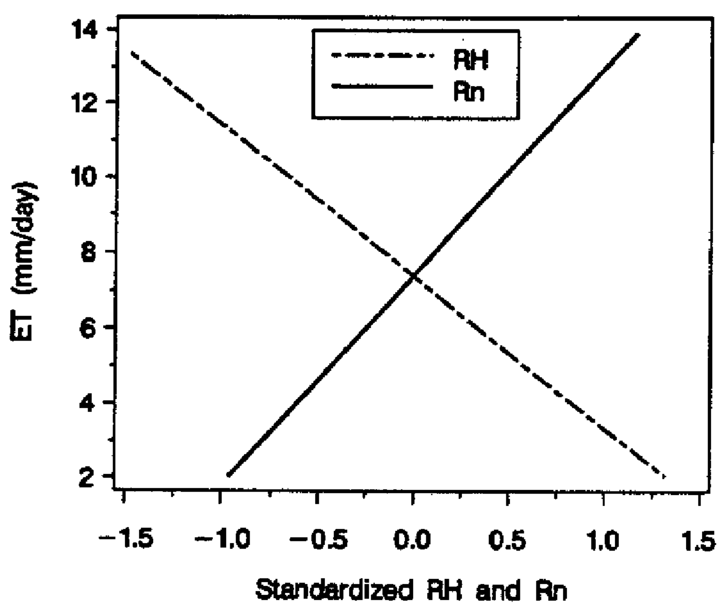

20m
65m
48m
10m windtower
Z
E
N
S
W
WX33
WX32
WX31
WX13
WX11
WX12
2.6m windtower
WX23
WX21
WX22
Cattail Fields
Dirt Road
Raingage
Water
Small Canal
Non-weighable Lysimeter
not to scale
Location of Weather Station
: Weather Sensor
: Anemometer for wind measurment

hahn-Figure 2



hahn-Figure 3





14
12
10
8
6
4
2
ET (mm/day)
RH
Rn
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
Standardized RH and Rn

hahn-Figure 5